DOCUMENT RESUME

ED 146 610 CS 203 730

AUTHOR Freedman, Sarah Warshauer
TITLE Models of the Evaluation Process.
PUB DATE 77
NOTE 8p.; Paper presented at the Annual Meeting of the Conference on College Composition and Communication (28th, Kansas City, Missouri, March 31-April 2, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Composition (Literary); English Instruction; *Evaluation Methods; *Evaluation Needs; *Grading; *Research Needs; Secondary Education; *Student Evaluation

ABSTRACT

This paper presents a model for evaluating student writing, that emphasizes the interplay between the paper and the evaluator. Past studies of composition evaluation have tended to focus on the paper or on the rater, but not both. The paper concludes that further experimental research on the evaluation process must be designed to examine the entire context of composition evaluation. (RL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED146610

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sarah Warshauer Freedman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND THE ERIC SYSTEM CONTRACTORS."

MODELS OF THE EVALUATION PROCESS

Revised Version of a Paper Presented at the Annual Conference on College Composition and Communication, Kansas City, 1977

By

SARAH WARSHAUER FREEDMAN

One pressing concern in English education today is student writing. Members of the National Council of Teachers of English (NCTE), the largest professional organization of language arts teachers in the United States, interpreted the latest results of the National Assessment of Educational Progress as "disturbing evidence that at least some of the best high school students are understimulated by the writing curriculum, and that there is wholesale stagnation throughout the high school years among the least able writers" (Mellon, 1975, p.29). They declared 1976 the year of composition and vowed to do "something" about the crisis in writing. Newsweek Magazine's December 8, 1975, cover story, "Why Johnny Can't Write," popularized the problem, stimulating public as well as professional concern.

The problems in the field of composition go far beyond whether or not students write more poorly now than they did in the past. Many teachers question their own ability and, for that matter, anyone's ability to teach composition. Furthermore, those convinced that they can and do teach students to improve their writing have difficulty demonstrating their convictions. The subjectivity of the composition evaluation process frequently

CS203730

results in erratic evaluations of student papers by teachers and by researchers. Joseph Williams in a keynote address at the annual meeting of of NCTE in 1976 emphasized how little researchers know about how and why evaluators or teachers judge student papers as they do. This inability to evaluate compositions has kept both pedagogy and research at a primitive level.

Most of the past attempts to solve problems in evaluation have either failed to answer or have ignored a central question: why do evaluators award the ratings they do to student papers? To answer this question researchers need to identify the factors in the papers, the raters, and the context of the evaluation that influence the evaluation.

Hiller, Marcotte, and Martin (1969) after completing a study of student essays suggested:

> if a given characteristic is present in an essay, does that characteristic affect the essay's quality as reflected in the grade assigned by expert graders. To answer this question we should have to manipulate the quality and quantity of relevant category items under an experimental Procedure. (p.274)

Most past research attempting to identify the factors influencing raters' judgments has been correlational and, therefore, could not establish patterns of influence. Experimental research is needed.

The correlational studies in composition evaluation, furthermore, have tended to focus on the paper or the rater, but not both. Paper-focused studies e.g. Page (1968), Slotnick and Knapp (1971), Hiller et all. (1969), Thompson (1976),

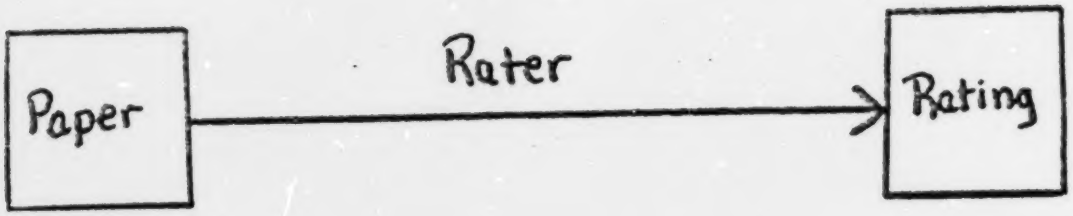
and Nold and Freedman (Forthcoming) correlated properties of the student papers such as the number of spelling errors and the length of essay with ratings of those papers. Rater-focused studies [e.g. Diederich et al. (1961) and Meyers et al. (1966)] correlated characteristics of the raters such as personal biases and degree of leniency with their ratings.

Part A of Figure 1 depicts the assumptions of the paper-focused studies. The rater is outside the bounds of the evaluation process because all raters are expected to give the same response, much like a computer. Part B of Figure 1 depicts the assumptions of the rater-focused studies. The paper remains outside the process. Disagreement between raters results from their biases and other personal qualities.

Figure 2 provides a more comprehensive model on which experimental studies could be based. Both the evaluator and the paper are central to the process. The reciprocal arrow connecting the rater and paper represents an interplay between the two. The qualities in the paper are meaningful in terms of how the evaluator perceives them. The evaluator creates part of the meaning of the paper, but qualities intrinsic to the paper set bounds on the meaning the evaluator may create. Most of the time the rating is based on this interplay. However, the line connecting the rating with the paper is dotted to allow for ratings that are not based totally on qualities within the paper. Rosner, for example, revealed that raters gave higher marks to papers they thought written by honors students than those they thought

PART A

Early Models of the Evaluation Process



PART B

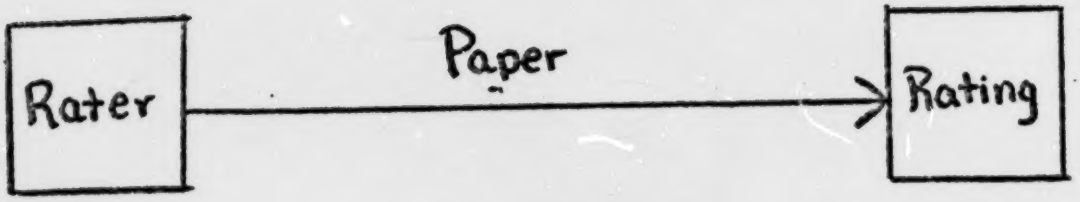


Figure 1

Present Model of the Evaluation Process

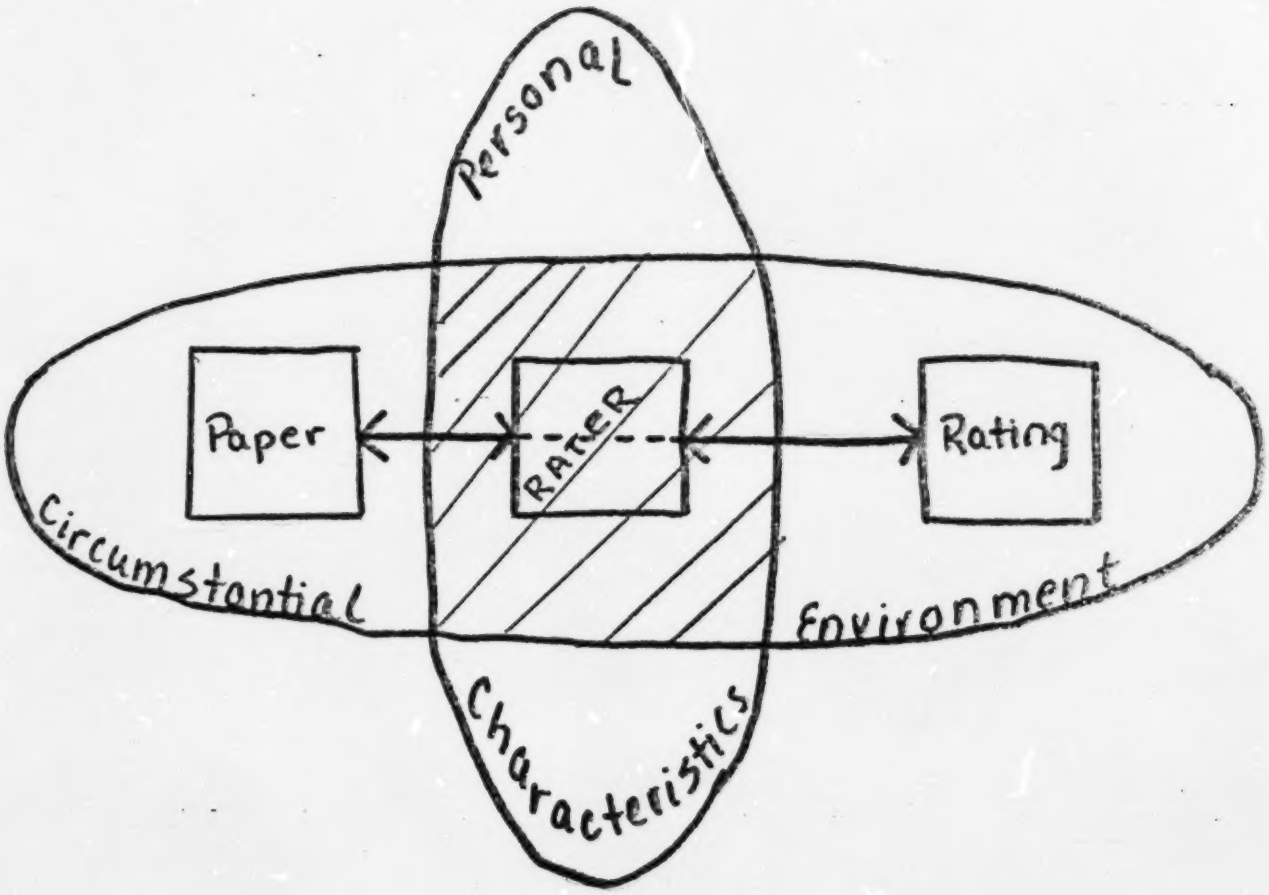


Figure 2

written by average students (in Diederich, 1974, pp. 11-13).

How the evaluator perceives and reacts to the paper is modulated by personal characteristics resulting from intrinsic qualities and past and present experiences as well as by a circumstantial environment manifest at the time of the reading. These personal characteristics likely to be important to the evaluation include: aptitudes, education, teaching and writing experience, prior knowledge about the writer, and personal anxieties and tensions. The circumstantial environment, the immediate context of the reading, consists of features like: training; time of day; length of task; physical environment--comfort of chairs, lighting of room, heating, noise level; the nature of the task--the kind of rating scale, types of papers to be read (narrative, expository), context of other papers, inferences about the writers based on reading the paper. The shaded area in Figure 2 illustrates the melding of these personal characteristics with the circumstantial environment during the evaluation process.

Experimental research on the evaluation process must be designed to allow examination of the entire context of the evaluation--a study of the interplay between the paper and the evaluator. I propose that the evaluator's qualities in conjunction with the paper itself determine his or her perception and evaluation of the paper. The evaluator's perception of the paper, in a particular context, written by a particular writer, under particular conditions, will determine the rating.

BIBLIOGRAPHY

Diederich, P. Measuring growth in English. Urbana, Ill.: National Council of Teachers of English, 1974.

_____, P., S. French, and S. Carlton. Factors in judgments writing ability. Research Bulletin 61-15. Princeton: Educational Testing Service, 1961.

Hiller, J., D. Marcotte, and T. Martin. Opinionation, vagueness, and specificity distinctions: essay traits measured by computer. American Education Research Journal, 1969, 6, 271-286.

Meyers, A., C. McConville, and W. Coffman. Simplex structure in the grading of essay tests. Educational and Psychological Measurement, 1966, 26, 41-54.

Nold, E. and S. Freedman. An analysis of readers' responses to essays. Research in the Teaching of English, forthcoming.

Page, E. Analyzing student essays by computer. International Review of Education, 1968, 14, 210-225.

Slotnick, H. and J. Knapp. Essay grading by computer: a laboratory phenomenon? English Journal, 1971, 60, 75-87.

Thompson, R. Predicting writing quality, writing weaknesses that dependably predict holistic evaluations of freshman compositions. English Studies Collection, Inc., 1976, 1.

Why Johnny Can't Write. Newsweek Magazine. December 8, 1975, pp. 58-65.

Williams, J. The significance of the written word in American culture. Paper presented at the annual meeting of the National Council of Teachers of English, Chicago, 1976.